A SAMPLE SELECTION APPROACH TO CENSORED DEMAND SYSTEMS

STEVEN T. YEN AND BIING-HWAN LIN

The multivariate sample selection model is extended to a nonlinear equation system with partial selection and applied to household meat consumption in China. Elasticity estimates differ from those obtained from conventional maximum likelihood and Tobit estimates. Chinese meat products are gross complements while net substitution also exists in some cases.

Key words: censoring, China, meat, sample selection, translog demand system.

The use of micro survey data has been popular in estimating consumer demand equations. Important features of microdata include censored dependent variables. Estimation procedures for censored consumer demand systems include the primal (Kuhn-Tucker) approach of Wales and Woodland (1983), dual (virtual-price) approach of Lee and Pitt (1986), and the Tobit system (Amemiya 1974) estimated by generalized maximum entropy (Golan, Perloff, and Shen 2001) and maximum simulated likelihood (Dong, Gould, and Kaiser 2004; Yen, Lin, and Smallwood 2003) procedures. Less efficient alternatives include the quasi-maximum likelihood estimator (Yen, Lin, and Smallwood 2003), the generalized method of moments estimator (Meyerhoefer, Ranney, and Sahn 2005), and a number of two-step estimators (Heien and Wessells 1990; Perali and Chavas 2000; Shonkwiler and Yen 1999). Yen (2005) recently proposed a maximum likelihood (ML) procedure for the multivariate sample selection model (MSSM), an extension of the bivariate sample selection model (Heckman 1979), which had motivated the procedures of Heien and Wessells (1990) and Shonkwiler and Yen (1999). The MSSM was developed in the context of linear equations and, in addition, is not strictly applicable for a partially selective equation system, viz., one in which only a subset of the equations is subject to sample selection. Because partial selection is the rule rather than exception in empirical applications, in this article the MSSM is extended to accommodate partial selection in a nonlinear equation system. The procedure is used to study meat consumption by urban households in China.

A Partially Selective Demand System

Let $\mathbf{x}$ be a vector of explanatory variables and $\boldsymbol{\theta}$ a vector of parameters, and consider a system of n demand equations in which each expenditure share w_i is generated by a deterministic function $f_i(\mathbf{x};\boldsymbol{\theta})$ and an unobservable error term v_i. The first k equations are subject to sample selection

$$(1)\qquad \begin{aligned} w_i &= d_i[f_i(\mathbf{x};\boldsymbol{\theta}) + v_i], \quad i = 1,\ldots,k \\ &= f_i(\mathbf{x};\boldsymbol{\theta}) + v_i, \quad i = k+1,\ldots n. \end{aligned}$$

Each indicator variable d_i is modeled with a binary probit

$$(2)\qquad d_i = 1(\mathbf{z}_i'\boldsymbol{\gamma}_i + u_i > 0), \quad i = 1,\ldots,k$$

where $1(\cdot)$ is a binary indicator function, $\mathbf{z}_i$ is a vector of variables, $\boldsymbol{\gamma}_i$ is a vector of parameters, and u_i is a random error.

The expenditure shares in equation (1) do not add up to unity unless $d_1 = \cdots = d_k = 1$. While adding-up can be accommodated in other ways, such as remapping of observed and censored variables (Wales and Woodland 1983; also see Dong, Gould, and Kaiser 2004), to limit the scope of the current article we take a simple approach of estimating the first $n-1$

Steven Yen is associate professor of Agricultural Economics at the University of Tennessee. Biing-Hwan Lin is agricultural economist at the Economic Research Service of the U.S. Department of Agriculture, Washington DC.

The authors thank two anonymous reviewers for their helpful comments. Research for this paper was supported by USDA-ERS Cooperative Agreement 43-3AEM-4-80052. The views in this paper are those of the authors and do not necessarily reflect the views or policies of the U.S. Department of Agriculture.

Amer. J. Agr. Econ. 88(3) (August 2006): 742–749

equations with the nth good treated as a residual category (cf., Pudney 1989, p. 155; also see Yen, Lin, and Smallwood 2003). The resulting ML estimates are not invariant with respect to the equation excluded, and this issue is addressed below by excluding alternative equations in estimation.

Denote $m = n - 1$ and, following Yen (2005), assume the concatenated error vector $[\mathbf{u}', \mathbf{v}']' \equiv [u_1, \ldots, u_k, v_1, \ldots, v_m]'$ is distributed as $(k+m)$-variate normal with zero mean and covariance matrix

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} = [\sigma_h \rho_{h\ell} \sigma_\ell] \tag{3}$$

such that Σ_{11} is $k \times k$, $\Sigma_{21} = \Sigma_{12}'$ is $m \times k$, Σ_{22} is $m \times m$, $\rho_{h\ell}$ are elements of a $(k+m) \times (k+m)$ error correlation matrix and σ_h are elements of a $(k+m)$-vector such that $\sigma_h = 1$ for $h = 1, \ldots, k$. For later use, define vectors $\mathbf{r} \equiv [r_1, \ldots, r_k]' \equiv [\mathbf{z}_i' \boldsymbol{\gamma}_i]$ and $\mathbf{v} \equiv [v_1, \ldots, v_m] \equiv [w_i - f_i(\mathbf{x}; \boldsymbol{\theta})]$.

In the likelihood construction below, the sample regimes are categorized by outcomes of $w_1, \ldots, w_k$ because $w_{k+1}, \ldots, w_m$ are not censored. Consider first a regime in which, without loss of generality, the first $\ell < k$ goods are censored such that

$$\begin{aligned} \mathbf{z}_i' \boldsymbol{\gamma}_i + u_i &\le 0, \quad i = 1, \ldots, \ell \\ \mathbf{z}_i' \boldsymbol{\gamma}_i + u_i &> 0, \quad i = \ell+1, \ldots, k \\ w_i = f_i(\mathbf{x}; \boldsymbol{\theta}) + v_i&, \quad i = \ell+1, \ldots, k, \\ & \qquad\quad k+1, \ldots, m. \end{aligned} \tag{4}$$

Let $\tilde{\mathbf{v}} \equiv [v_{\ell+1}, \ldots, v_m]$ be an $(m-\ell)$-vector with the first ℓ elements of $\mathbf{v}$ excluded. Then, $[\mathbf{u}', \tilde{\mathbf{v}}']'$ is $(k+m-\ell)$-variate normal with zero mean and covariance matrix $\tilde{\Sigma}$, where $\tilde{\Sigma}$ is a $(k+m-\ell) \times (k+m-\ell)$ submatrix of Σ in equation (3) with rows and columns corresponding to $v_1, \ldots, v_\ell$ excluded. Partition $\tilde{\Sigma}$ at the kth row and column

$$\tilde{\Sigma} = \begin{bmatrix} \Sigma_{11} & \tilde{\Sigma}_{12} \\ \tilde{\Sigma}_{21} & \tilde{\Sigma}_{22} \end{bmatrix} \tag{5}$$

such that Σ_{11} is $k \times k$, $\tilde{\Sigma}_{21} = \tilde{\Sigma}_{12}'$ is $(m-\ell) \times k$ and $\tilde{\Sigma}_{22}$ is $(m-\ell) \times (m-\ell)$. Let $g(\tilde{\mathbf{v}})$ be the marginal probability density function (pdf) of $\tilde{\mathbf{v}} \sim N(0, \tilde{\Sigma}_{22})$ and $h(\mathbf{u} \mid \tilde{\mathbf{v}})$ the conditional pdf of $\mathbf{u} \mid \tilde{\mathbf{v}} \sim N(\boldsymbol{\mu}_{\mathbf{u}|\tilde{\mathbf{v}}}, \Sigma_{\mathbf{u}|\tilde{\mathbf{v}}})$, where

$$\boldsymbol{\mu}_{\mathbf{u}|\tilde{\mathbf{v}}} = \tilde{\Sigma}_{12} \tilde{\Sigma}_{22}^{-1} \tilde{\mathbf{v}} \tag{6}$$

$$\Sigma_{\mathbf{u}|\tilde{\mathbf{v}}} = \Sigma_{11} - \tilde{\Sigma}_{12} \tilde{\Sigma}_{22}^{-1} \tilde{\Sigma}_{21}. \tag{7}$$

See Kotz, Johnson, and Balakrishnan (2000) for conditional moments of the multivariate normal distribution. Define a diagonal matrix $\mathbf{D} = \mathrm{diag}[2d_1 - 1, \ldots, 2d_k - 1]$ and, finally, denote $\Phi_q(\mathbf{a}; \mathbf{b})$ as q-variate normal cumulative distribution function (cdf) with zero means, covariance $\mathbf{b}$ and finite upper integration limit $\mathbf{a}$. Then, the likelihood contribution for this regime is

$$\begin{aligned} L &= g(\tilde{\mathbf{v}}) \int_{-\infty}^{-\mathbf{z}'_1 \boldsymbol{\gamma}_1} \cdots \int_{-\infty}^{-\mathbf{z}'_\ell \boldsymbol{\gamma}_\ell} \int_{-\mathbf{z}'_{\ell+1} \boldsymbol{\gamma}_{\ell+1}}^{\infty} \cdots \int_{-\mathbf{z}'_k \boldsymbol{\gamma}_k}^{\infty} \\ &\quad \times h(u_1, \ldots, u_\ell, u_{\ell+1}, \ldots, u_k \mid \tilde{\mathbf{v}}) \\ &\quad \times du_k \cdots du_{\ell+1} du_\ell \cdots du_1 \\ &= g(\tilde{\mathbf{v}}) \int_{\mathbf{u}^* \le \mathbf{D}\mathbf{r}} h(\mathbf{u}^* \mid \tilde{\mathbf{v}}) \, d\mathbf{u}^* \\ &= g(\tilde{\mathbf{v}}) \Phi_k(\mathbf{D}(\mathbf{r} + \boldsymbol{\mu}_{\mathbf{u}|\tilde{\mathbf{v}}}); \mathbf{D}' \Sigma_{\mathbf{u}|\tilde{\mathbf{v}}} \mathbf{D}) \end{aligned} \tag{8}$$

where conditional moments $\boldsymbol{\mu}_{\mathbf{u}\,|\,\tilde{\mathbf{v}}}$ and $\Sigma_{\mathbf{u}|\tilde{\mathbf{v}}}$ are defined in equations (6) and (7).

Consider next a regime in which the first k goods are censored, characterized by

$$\begin{aligned} \mathbf{z}_i' \boldsymbol{\gamma}_i + u_i &\le 0, \quad i = 1, \ldots, k \\ w_i = f_i(\mathbf{x}; \boldsymbol{\theta}) + v_i&, \quad i = k+1, \ldots, m. \end{aligned} \tag{9}$$

This is a special case of the first regime in equation (4) when $\ell = k$, for which the likelihood contribution is also equation (8) but constructed with error vector $\tilde{\mathbf{v}} \equiv [v_{k+1}, \ldots, v_m]$ and covariance matrices $\tilde{\Sigma}_{21}$, $\tilde{\Sigma}_{12}$, and $\tilde{\Sigma}_{22}$ in which rows and columns corresponding to $v_1, \ldots, v_k$ are excluded.

The last regime is one in which $w_1, \ldots, w_k$ are all positive, characterized by

$$\begin{aligned} \mathbf{z}'_i \boldsymbol{\gamma}_i + u_i &> 0, \quad i = 1, \ldots, k \\ w_i = f_i(\mathbf{x}; \boldsymbol{\theta}) + v_i&, \quad i = 1, \ldots, k, \\ & \qquad\quad k+1, \ldots, m. \end{aligned} \tag{10}$$

This is another special case of the first regime in equation (4) where $\ell = 0$, for which the likelihood contribution (equation (8)) is constructed with error vector $\tilde{\mathbf{v}} = \mathbf{v}$ and covariance matrices $\tilde{\Sigma}_{21} = \tilde{\Sigma}_{12}' = \Sigma_{21}$ and $\tilde{\Sigma}_{22} = \Sigma_{22}$.

The MSSM of Yen (2005, equation (1)) corresponds to the above when $k = m$ and Σ_{11} is $m \times m$. For this fully selective model the likelihood contributions for all three regimes can be

Table 1. Variable Definitions and Sample Statistics (Sample Size 2,827)

Variable	Mean	SD
Quantities (kg. per person per annum)		
Beef (mutton included)	3.54	4.31
Consuming households (90.3% of sample)	3.91	4.37
Pork	9.39	7.70
Poultry	18.19	15.47
Fish	17.73	10.69
Expenditures (Yuan per person per annum)		
Beef	50.60	59.74
Consuming households	56.01	60.39
Pork	212.52	136.22
Poultry	154.99	140.94
Fish	268.64	293.24
Prices (Yuan/kg.)		
Beef	14.57	3.74
Pork	15.65	4.54
Poultry	13.15	6.32
Fish	11.96	2.34
Age (of household head)	48.66	11.61
Size (of household)	3.09	0.82
Dummy variables (1 = yes; 0 otherwise)		
College	0.22	
High school (junior/senior, tech. school)	0.70	
Less than high school (reference)	0.08	
Region 1 (Beijing)	0.17	
Region 2 (Tianjin, Hebei)	0.05	
Region 3 (Liaoning, Jelin, Heilongjiang)	0.11	
Region 4 (Shanghai)	0.17	
Region 5 (Anhuei, Shandong, Henan)	0.10	
Region 6 (Fujian, Jiangxi)	0.07	
Region 7 (Hubei, Hunan)	0.07	
Region 8 (Guangzhou)	0.10	
Region 9 (Guangdong, Hainan)	0.05	
Region 10 (Guangxi)	0.04	
Region 11 (Jiangsu, Jhejiang) (reference)	0.07	

Source: Urban Household Survey, National Statistical Bureau, 2000.

adopted from equation (8) except the second (all-zero) regime, for which the terms $g(\tilde{\mathbf{v}})$ and $\boldsymbol{\mu}_{\mathbf{u}|\tilde{\mathbf{v}}}$ would be removed and $\boldsymbol{\Sigma}_{\mathbf{u}|\tilde{\mathbf{v}}}$ replaced with the unconditional covariance matrix $\boldsymbol{\Sigma}_{11}$ (Yen 2005, equation (6)).

To demonstrate the proposed estimator we use the translog demand system (Christensen, Jorgenson, and Lau 1975), with deterministic shares

$$(11)\quad f_i(\mathbf{x};\boldsymbol{\theta}) = \frac{\alpha_i + \sum_{j=1}^{n} \beta_{ij} \log p_j}{-1 + \sum_{k=1}^{n} \sum_{j=1}^{n} \beta_{kj} \log p_j}, \quad i = 1, \ldots, n$$

where p_j are expenditure-standardized prices and α_i and β_{ij} are parameters. Homogeneity follows from use of the standardized prices, and symmetry ($\beta_{ij} = \beta_{ji} \forall i, j$) is imposed. Demographic variables s_j are incorporated in equation (11) by parameterizing α_i such that $\alpha_i = \alpha_{i0} + \sum_j \alpha_{ij} s_j, i = 1, \ldots, m$.[1]

For products that are censored, elasticities are calculated from the unconditional means of the expenditure shares. Based on the bivariate normality of errors $[u_i, v_i]'$ for each i, the unconditional means of w_i are

$$(12)\quad E(w_i) = \Phi(\mathbf{z}_i'\boldsymbol{\gamma}_i)\, f_i(\mathbf{x};\boldsymbol{\theta}) + \sigma_{k+i,i}\phi(\mathbf{z}_i'\boldsymbol{\gamma}_i), \quad i = 1, \ldots, k$$

[1] Usual restrictions are not imposed on α_{i0} and α_{ij} as in a conventional translog demand system because adding-up does not hold even with these restrictions and are accommodated in another way, as described above.

Table 2. ML Estimates of Multivariate Sample Selection Model: Translog Demand System

	Beef		Pork		Poultry	
Variables	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Selection equations (γ_{ij})						
Constant	−0.258	0.216				
Age	0.043*	0.026				
Size	0.009	0.033				
College	−0.139	0.107				
High school	0.028	0.097				
Region 1	1.869***	0.149				
Region 2	1.745***	0.249				
Region 3	1.586***	0.138				
Region 4	0.671***	0.093				
Region 5	0.746***	0.094				
Region 6	0.047	0.111				
Region 7	0.672***	0.103				
Region 8	1.094***	0.115				
Region 9	0.103	0.112				
Region 10	1.171***	0.120				
Demand system: demographic variables (α_{ij})						
Constant	−0.239***	0.023	−0.568***	0.034	−0.177***	0.028
Age	0.001	0.002	−0.015***	0.003	0.010***	0.002
Size	0.003	0.002	0.010***	0.003	0.003	0.003
Region 1	−0.068***	0.010	0.011	0.009	0.002	0.006
Region 4	0.042***	0.008	0.093***	0.016	−0.017*	0.007
Region 8	0.055***	0.008	0.132***	0.020	−0.086***	0.015
Quadratic price terms (β_{ij})						
Beef	−0.077***	0.011				
Pork	−0.004	0.009	−0.265***	0.033		
Poultry	0.033***	0.007	0.062***	0.015	−0.101***	0.014
Fish	0.022***	0.007	0.156***	0.024	0.024***	0.007
Std. dev. (σ_h)	0.097***	0.003	0.137***	0.002	0.110***	0.002
Log-likelihood		5,532.134				

Note: Triple (***) and single (*) asterisks indicate significance at the 1% and 10% levels, respectively. The coefficient of the quadratic log-price term (β_{44}, where 4 indicates fish), not reported due to space consideration, is −0.144 and has a standard error of 0.016.

where $\phi(\cdot)$ and $\Phi(\cdot)$ are univariate standard normal pdf and cdf, and $\sigma_{k+i,i}$ is error covariance between u_i and v_i. Differentiation of equation (12) gives demand elasticities for the first k goods (Yen, Kan, and Su 2002), elasticities for goods $k+1,\ldots,m$ can be derived from the conventional translog demand system, and elasticities for the nth goods by using the adding up restriction (Yen, Lin, and Smallwood 2003).[2]

Data and Application

Data for household meat consumption were compiled from the 2000 Urban Household Survey collected by China's National Statistical Bureau. Households from the West and pastoral regions are excluded because consumption patterns in these regions are likely to be very different from the rest of the country.

Table 3. ML Estimates of Error Correlation Coefficients

Share Equation	Selection Equation Beef (u_1)	Share Equation Beef (v_1)	Share Equation Pork (v_2)
Beef (v_1)	−0.539 (0.030)		
Pork (v_2)	−0.282 (0.030)	−0.251 (0.026)	
Poultry (v_3)	−0.139 (0.031)	−0.186 (0.019)	−0.390 (0.017)

Note: Asymptotic standard errors are in parentheses. All parameter estimates are significant at the 1% level.

[2] For censored product i the probability of a positive observation is $\Pr(w_i > 0) = \Phi(\mathbf{z}_i'\boldsymbol{\gamma}_i)$ and the conditional mean of expenditure share is $E(w_i \mid w_i > 0) = f_i(\mathbf{x};\boldsymbol{\theta}) + \sigma_{k+i,i}\phi(\mathbf{z}_i'\boldsymbol{\gamma}_i)/\Phi(\mathbf{z}_i'\boldsymbol{\gamma}_i)$. Additional elasticities can be derived by differentiating these expressions (Yen, Kan, and Su 2002).

Table 4. Demand Elasticities

Product	Price of Beef	Price of Pork	Price of Poultry	Price of Fish	Total Expenditure
Uncompensated elasticities					
Beef	−0.33***	−0.03	−0.28***	−0.13***	0.77***
	(0.06)	(0.08)	(0.06)	(0.05)	(0.04)
Pork	−0.02	−0.28***	−0.16***	−0.40***	0.85***
	(0.02)	(0.05)	(0.03)	(0.03)	(0.02)
Poultry	−0.18***	−0.34***	−0.51***	−0.05*	1.08***
	(0.03)	(0.05)	(0.04)	(0.03)	(0.03)
Fish	−0.05***	−0.51***	−0.07***	−0.53***	1.16***
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
Compensated elasticities					
Beef	−0.23***	0.24***	−0.12**	0.11**	
	(0.06)	(0.08)	(0.06)	(0.05)	
Pork	0.09***	0.02	0.02	−0.13***	
	(0.02)	(0.05)	(0.03)	(0.03)	
Poultry	−0.04	0.04	−0.28***	0.28***	
	(0.03)	(0.05)	(0.04)	(0.03)	
Fish	0.10***	−0.10***	0.18***	−0.17***	
	(0.02)	(0.03)	(0.02)	(0.03)	

Note: Asymptotic standard errors are in parentheses. Triple (***), double (**), and single (*) asterisks indicate significance at the 1%, 5%, and 10% levels, respectively.

To limit the scope of the analysis, this study focuses on four of the more popular meat products: beef (mutton included), pork, poultry, and fish. Meats from large animals such as horse, mule, and donkey are excluded.

From the reported expenditure and quantity of each food item consumed, price was derived as the unit value, and missing prices for nonconsuming households were replaced with regional averages. While this zero-order imputation is commonly used in empirical applications (e.g., Dong, Gould, and Kaiser 2004; Yen, Lin, and Smallwood 2003), further applications might address this missing-price issue. Demographic variables used include household size, age of the household head, and dummy variables indicating education of the household head, along with ten other dummy variables for regions.

Our sample consists of 2,827 urban households. Pork, poultry, and fish are consumed by nearly all (over 99%) households in the sample, while about 90.3% of the sample consume beef during the year.[3] The large proportion of consuming households for beef and near-absence of zeros in the other meat products are due to the long duration of the survey (one year). Definitions of regions and all other variables, as well as their sample statistics, are presented in table 1.

ML estimation is carried out with the fish equation excluded from the system. Drawing on the discrete random utility theory of Pudney (1989, pp. 160–2), which motivates a similar selectivity model, specifically the double-hurdle model, only demographic (nonprice) variables are included in the selection equation. The ML estimates, along with their robust standard errors (White, 1982), are presented in tables 2 and 3. Age is significant at the 10% level and eight of the ten regional variables are significant at the 1% level in the selection equation for beef. Age and household size are also included in the share equations, along with three dummy variables for the large cities of Beijing, Shanghai, and Guangzhou, and statistical significance justifies the use of these variables. All but one of the quadratic price coefficients (β_{ij}) are significant at the 1% level. Estimates for error standard deviations of all share equations are significant at the 1% level of significance (table 2), as are all error correlation coefficients (table 3). Significance of the error correlations suggests that correction for sample selection is important despite the small proportion of zero observations for

[3] Proportions of consuming households are higher for beef (95.7%) and lower for pork (86.8%), poultry (88.8%), and fish (91.5%) for the pastoral region. Thus, regional differences may masquerade as censoring in the national sample—another reason to exclude the pastoral region from the sample.

beef. The selection equation correctly predicts the binary outcomes for beef 79% of the time, at a probability cutoff of 0.5 (Wooldridge 2002, p. 465). Predicted shares lie outside the unit circle for only 21 observations for beef and 1 observation for fish. All predictions for pork and poultry lie within the unit circle. The correlation coefficients between observed and predicted expenditure shares are 0.59 for beef, 0.62 for pork, 0.37 for poultry, and 0.72 for fish.

Table 4 presents the demand elasticities and their standard errors, calculated by the *delta method* (Spanos 1999). All uncompensated own-price elasticities are negative, well below unity, and significant at the 1% level. All uncompensated cross-price elasticities are significant (except between pork and beef) and negative at the 10% level or lower, suggesting gross complementarity among the meat products. Expenditure elasticities are below unity for beef and pork but above unity for poultry and fish. Unlike the uniformly negative uncompensated cross-price effects, the compensated elasticities suggest net substitution between fish and beef and between fish and poultry, and net complementarity between fish and pork and between poultry and beef. All compensated own-price elasticities are negative and significant except pork but, due to the positive expenditure elasticities, are smaller in absolute values than their uncompensated counterparts.

To investigate the invariance issue mentioned above, the demand share equations are estimated by excluding alternative equations from the system. The results, presented in table 5, suggest that own-price and expenditure elasticities are robust regardless of whether fish, poultry, or pork is excluded.[4] However, when beef is excluded, ML estimation produces very different own- and expenditure elasticities for beef but similar elasticities for the other products. Note that when beef is omitted, estimation is carried out with conventional ML procedure (because the other equations are nearly noncensored) and that the results are invariant with respect to the equation omitted. Thus, the different elasticities for beef are likely the results of ignoring its zero observations and not of omission of the equation per se. Also presented in table 5 are own-price and expenditure elasticities from the Tobit system estimates (see Yen, Lin, and Smallwood 2003), whose elasticities

[4] Cross-price elasticities, not presented, are also robust with respect to equations omitted.

Table 5. Elasticities from Excluding Alternative Equations and Comparisons with Conventional and Tobit Estimates

	Equation Omitted								Tobit Estimates Omitted-Fish	
	Fish		Poultry		Pork		Beef			
Product	Own-Price	Expenditure	Own-Price	Expenditure	Own-Price	Expenditure	Own-Price	Expenditure	Own-Price	Expenditure
Beef	−0.33	0.77	−0.33	0.72	−0.28	0.75	−0.19†	0.94	−0.42	1.05
	(0.06)	(0.04)	(0.06)	(0.04)	(0.06)	(0.04)	(0.13)	(0.04)	(0.09)	(0.34)
Pork	−0.28	0.85	−0.27	0.91	−0.25	0.84	−0.27	0.81	−0.28	0.81
	(0.05)	(0.02)	(0.06)	(0.05)	(0.05)	(0.02)	(0.06)	(0.02)	(0.04)	(0.15)
Poultry	−0.51	1.08	−0.61	1.09	−0.50	1.08	−0.50	1.06	−0.51	1.07
	(0.04)	(0.03)	(0.05)	(0.03)	(0.04)	(0.03)	(0.04)	(0.02)	(0.04)	(0.15)
Fish	−0.53	1.16	−0.45	1.11	−0.51	1.17	−0.50	1.17	−0.51	1.13
	(0.02)	(0.02)	(0.03)	(0.05)	(0.03)	(0.02)	(0.02)	(0.01)	(0.02)	(0.13)

Note: Asymptotic standard errors are in parentheses. Dagger (†) indicates insignificance at the 10% significance level. All other elasticities are significant at the 1% level. Estimation with beef omitted produces conventional ML estimates (no censoring).

for beef are very different from the MSSM results. To sum up, lack of invariance in the parameter estimates does not appear to be a discernable issue for the current application, whereas conventional and Tobit estimation procedure produce very different demand elasticities from the MSSM estimates.

Concluding Remarks

With the growing popularity of microdata in empirical analysis, interest in censored data issues has continued to grow. This article contributes to the censored demand system literature by extending the MSSM to one with nonlinear equations and partial selection. While it is possible to construct a model with mixed selection, Tobit and continuous mechanisms, a system procedure to address other causes of zeros such as conscientious abstention and infrequency of purchases (Pudney 1988) is yet to be developed. The difficulty may lie in the fact that the causes of zeros are often mixed and not clear-cut. For this reason the sample selection approach proposed in this article is powerful for being able to accommodate selectivity (zeros) of unknown causes.

For the current application, zero observations occur in one equation and estimation requires only evaluation of univariate cdf's. For a larger system with many censored equations, the multiple probability integrals would have to be evaluated with existing simulation techniques (Hajivassiliou 1993). For these large systems a two-step estimator (e.g., Shonkwiler and Yen 1999), though statistically inefficient, avoids the computational complexity of the ML estimators and remains an attractive alternative. For a system where only a subset of equations is censored, the partially selective system proposed here formally motivates two-step estimation of the system whereby, as intuition might interestingly suggest, selectivity terms would be included only for equations with zero observations. This in fact was the procedure, which produced the two-step results in Yen, Lin, and Smallwood (2003).

[Received May 2004; accepted October 2005.]

References

Amemiya, T. 1974. "Multivariate Regression and Simultaneous Equation Models when the Dependent Variables Are Truncated Normal." *Econometrica* 42(6):999–1012.

Christensen, L.R., D.W. Jorgenson, and L.J. Lau. 1975. "Transcendental Logarithmic Utility Functions." *American Economic Review* 65(3):367–83.

Dong, D., B.W. Gould, and H.M. Kaiser. 2004. "Food Demand in Mexico: An Application of the Amemiya-Tobin Approach to the Estimation of a Censored System." *American Journal of Agricultural Economics* 86(4):1094–107.

Golan, A., J.M. Perloff, and E.Z. Shen. 2001. "Estimating a Demand System with Nonnegativity Constraints: Mexican Meat Demand." *Review of Economics and Statistics* 83(3):541–50.

Hajivassiliou, V.A. 1993. "Simulation Estimation Methods for Limited Dependent Variable Models." *Handbook of Statistics*, vol. 11. G.S. Maddala, C.R. Rao, and H.D. Vinod, eds., pp. 519–43. Amsterdam: North-Holland.

Heckman, J.J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1):153–61.

Heien, D., and C.R. Wessells. 1990. "Demand Systems Estimation with Microdata: A Censored Regression Approach." *Journal of Business and Economic Statistics* 8(3):365–71.

Kotz, S., N.L. Johnson, and N. Balakrishnan. 2000. *Continuous Multivariate Distributions, vol. 1: Models and Applications*, 2nd ed. New York: John Wiley & Sons.

Lee, L-F., and M.M. Pitt. 1986. "Microeconometric Demand Systems with Binding Nonnegativity Constraints: The Dual Approach." *Econometrica* 54(5):1237–42.

Meyerhoefer, C.D., C.K. Ranney, and D.E. Sahn. 2005. "Consistent Estimation of Censored Demand Systems Using Panel Data." *American Journal of Agricultural Economics* 87(3):660–72.

Perali, F., and J.P. Chavas. 2000. "Estimation of Censored Demand Equations from Large Cross-Section Data." *American Journal of Agricultural Economics* 82(4):1022–37.

Pudney, S.E. 1988. "Estimating Engel Curves: A Generalization of the P-Tobit Model." *Finish Economic Papers* 1(2):129–47.

——. 1989. *Modelling Individual Choice: The Econometrics of Corners, Kinks, and Holes*. Cambridge, UK: Blackwell Publishers.

Shonkwiler, J.S., and S.T. Yen. 1999. "Two-Step Estimation of a Censored System of Equations." *American Journal of Agricultural Economics* 81(4):972–82.

Spanos, A. 1999. *Probability Theory and Statistical Inference: Econometric Modeling with Observational Data*. Cambridge, UK: Cambridge University Press.

Wales, T.J., and A.D. Woodland. 1983. "Estimation of Consumer Demand Systems with Binding Non-Negativity Constraints." *Journal of Econometrics* 21(3):263–85.

White, H. 1982. "Maximum Likelihood Estimation of Misspecified Models." *Econometrica* 50(1):1–25.

Wooldridge, J.M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Yen, S.T. 2005. "A Multivariate Sample-Selection Model: Estimating Cigarette and Alcohol Demands with Zero Observations." *American Journal of Agricultural Economics* 87(2):453–66.

Yen, S.T., K. Kan, and S. Su. 2002. "Household Demand for Fats and Oils: Two-Step Estimation of a Censored Demand System." *Applied Economics* 34(14):1799–806.

Yen, S.T., B. Lin, and D.M. Smallwood. 2003. "Quasi and Simulated Likelihood Approaches to Censored Demand Systems: Food Consumption by Food Stamp Recipients in the United States." *American Journal of Agricultural Economics* 85(2):458–78.